TY FOR THE PSYCHOLOGICAL STUDY OF SOCIAL ISSUES
(SPSSI)

sory Committee

ampbell, *Northwestern University*
, *New York University*
ı, *New York University*
is, *National Institute of Mental Health*
ds, *University of Michigan*
ɜ, *Stanford University*
ɔkins, *Columbia University*
ı, *Brooklyn College of the City University of New York*
Johns Hopkins University
Smith, *University of California, Santa Cruz*
bé, *University of Oklahoma*

n 9 of the American Psychological Association, SPSSI began as and be a unique organization. From its very beginning in 1936 it s the nexus of its concern the application of behavioral science ıe major social dilemmas of modern man; war, poverty, intergroup ban stress, industrial strife, and other problems. Implicit in this he compelling assumption that science could not be value free; that to ameliorate the ills of modern society its systematic research uld have political and social consequences as well as scientific ones. ght when he stated that research should be socially useful as well as meaningful, but this could only occur if the researchers involved d, organized, and directed their investigations with this end in

Excerpt from 1972 Presidential Message by
Harold M. Proshansky

Volume 2

Handbook of Evaluation Research

Edited by

MARCIA GUTTENTAG
Harvard University

and

ELMER L. STRUENING
Columbia University

SAGE Publications Beverly Hills • London

For information address:

SAGE PUBLICATIONS, INC.
275 South Beverly Drive
Beverly Hills, California 90212

SAGE PUBLICATIONS LTD
St George's House / 44 Hatton Garden
London EC1N 8ER

Printed in the United States of America

International Standard Book Number 0-8039-0429-0

Library of Congress Catalog Card No. 74-15764

FIRST PRINTING

CONTENTS

5

COST-EFFECTIVENESS ANALYSIS IN EVALUATION RESEARCH

HENRY M. LEVIN

Stanford University

I. INTRODUCTION

Costs and Public Policy Choices

The purpose of evaluation research is to obtain information that might be used to choose among alternative policies or programs for achieving social objectives. This decision orientation suggests that attempts to ascertain the impacts of the various approaches are not adequate in themselves to make such choices. Associated with any alternative is not only an impact or effect, but also a sacrifice or cost. The lower the cost for obtaining a given result, the greater will be the total impact of the social resources devoted to the problem. The focus of cost-effectiveness analysis in evaluation research is to determine that strategy or combination of strategies that maximizes the desired result for any particular resource or budget constraint.

In order to understand this concern more fully, it is necessary to consider the fact that government agencies and other institutions are faced with finite budgets and other resources for achieving their objectives.[1] Although each agency may have a relatively narrow set of goals, such as reducing crime, improving educational results, curbing pollution, improving nutrition, reducing infant mortality, and so on, there are presumably alternative methods for accomplishing these tasks. Traditionally, evaluation research has occupied itself only with comparing these alternatives with respect to their results. The costs of alternatives have not been considered. Yet, it is obvious that the less the cost of obtaining any particular set of

AUTHOR'S NOTE: This paper was prepared as a chapter for the *Handbook of Evaluation Research* being produced by The Society for the Psychological Study of Social Issues under the editorship of Marcia Guttentag. The author wishes to acknowledge the assistance of Linda Gunzel and Maureen McNulty in the preparation of the manuscript.

results, the greater will be the contribution of the program toward achieving agency goals because the limited resources will provide a greater impact.

In most respects, cost-effectiveness analysis is not a new form of evaluation research so much as it is one that attempts to integrate cost considerations into standard evaluation research designs. Because a major portion of the cost-effectiveness analysis is based upon the comparative effectiveness of the various alternatives, the approach presumes that such information will be provided by experimental or quasi-experimental research. This information on the probable differences in outcomes of particular strategies is then combined with data on the costs of implementing them in order to make cost-effectiveness comparisons. From these comparisons one can make tentative recommendations among alternatives on the basis of which approaches will maximize the desired outcomes for any particular level of resource use.

The vast majority of evaluation research endeavors fail to consider the cost component. This neglect appears to be largely attributable to the fact that much of the methodological basis for evaluation derives from the experimental sciences and particularly their applications in psychology. Because the preoccupation of such studies is with ascertaining whether there are "statistically significant" differences among experimental treatments and between experimental and control populations, this focus has also dominated much evaluation research. Questions of whether statistically significant differences in outcomes are socially significant, and cost analyses of such differences, have not been an important part of the evaluation agenda until quite recently.

Perhaps the willingness to nominate for policy consideration any program that shows statistically significant results in outcomes over alternatives derives from the fact that only too rarely can such differences be found. Under these circumstances, any differences are thought to be important and worthy of being used for policy recommendations by researchers. But such zeal may provide a very misleading answer once costs are taken into account. That is, the alternative that appears to yield better results in terms of comparative effectiveness may have costs that far outweigh its superiority in results.

This natural tendency to ignore costs is reflected by the following example. A study of computer-assisted instruction in mathematics found that children in grades 1-6 who had received such instruction showed greater gains on standardized tests than did students in matched control groups (Suppes and Morningstar, 1969a, 1969b). While this finding is certainly of interest to educational researchers and policy makers, it lacks a comparison of relative costs. Can the same gain in mathematics achievement be obtained for less cost by using a modification of the traditional mode of instruction?

The answer to that question is implied by other data collected for the experiment. It was found that one of the matched "control schools" using conventional instruction showed greater test score gains for grades 4 and 5 than did its computer-assisted counterpart. Further inquiry revealed that the teachers and administrators at the control school had instituted an additional twenty-five minutes per day of classroom instruction and practice in arithmetic for those two grades. The authors concluded tentatively "that twenty-five extra minutes of classroom drill can be

more beneficial than five to eight minutes per day of computer-based drill" (Suppes and Morningstar, 1969b: 19).

While precise cost data are not readily available, estimates of comparative costs can be obtained. Because five hours per day of teacher instruction cost the schools in the sample about $350 per child for the year, an additional twenty-five minutes of such instruction had an added cost of about $35 per year. It appeared that the cost of a fully utilized and efficient system of computer-assisted instruction for five to eight minutes of drill and practice a day would be on the order of about $150 per year.[2] Thus, the computer-assisted alternative was about four times as costly as traditional instruction for achieving the same objective. Because the schools were already spending about $700 annually for each pupil, the computer-augmented approach would have required a 20% addition to the school budget in contrast with only a 5% increase for using conventional methods to increase mathematics performance of students.

In the world of public policy, the costs as well as the effects of alternative strategies must be considered in order to maximize the impact of the social resources devoted to the public good. The purpose of this exposition is to provide a framework for integrating costs into the evaluation framework in order to carry out a cost-effectiveness analysis of the various policy choices. Emphasis will be placed on the conceptual framework, methodology, and examples. As in most areas of evaluation, there is no standard set of applied rules that can be mastered to carry out cost-effectiveness comparisons. Rather there is a set of principles and considerations that must be combined with sensitivity, ingenuity, and intuition on the part of the analyst. Much of the focus of this description will be on the principles of their application, with the hope that these have some transferability to other policy evaluations under other circumstances and in other domains.

Cost-Effectiveness, Cost-Benefit, and Cost-Utility Analysis

Before proceeding with a review of the concepts and their implementations, it is useful to review the history of cost-effectiveness analysis and to compare it with its close relatives, cost-benefit and cost-utility analysis. The father of cost-effectiveness analysis is the more general cost-benefit analysis. The cost-benefit framework allows one to compare the costs and benefits to society of the various policy alternatives confronted. In its most refined form, the technique can be used to take account of such complexities as alternatives characterized by differences in the time allocation of benefits and costs and differences in who receives the benefits and who pays the costs (e.g., rich vs. poor or young vs. old).[4] Moreover, since there is an attempt to compare the monetary value of benefits with the monetary value of costs, the cost-benefit calculus enables the evaluator to use a common yardstick to assess the relative attractiveness of alternatives. Thus, by calculating the costs and benefits of policy alternatives in terms of monetary values, one can compare such dimensions as rates of return on investment, net differences between costs and benefits (net present values), and benefit-to-cost ratios. Of course no alternative would be undertaken whose costs exceeded benefits, and in general the ones that would be selected would be those that maximized the total social benefits relative to costs.

The earliest and perhaps the most imaginative use of cost-benefit analysis for

public decision making has been in the evaluation of water resource projects. The Flood Control Act of 1936 required that the Corps of Army Engineers certify a project as feasible unless its total benefits exceeded its total costs. Accordingly, the last three decades have seen extensive development of techniques for measuring the benefits and costs of public investment in activities designed to improve navigation and flood control; to increase the output of hydroelectric power and water usable for agricultural, industrial, and domestic uses; and to expand and enhance the recreational potentiality of water resources (Eckstein, 1958; Krutilla and Eckstein, 1958).

In the application of cost-benefit analyses to water resource projects, most of the costs and benefits can be measured in monetary terms. For example, the costs of dredging and the construction of dams and other facilities, as well as costs of maintenance and operation, can generally be calculated directly from engineering data and cost experiences with similar projects. The monetary values of such benefits as hydroelectric power and water for agricultural, industrial, and domestic purposes can also be estimated by using the market prices of these outputs. The value of flood control can be assessed by deriving estimates of property destruction and loss of life that will be avoided by taming rampant waters. While recreational benefits are harder to determine, several techniques exist for calculating these (Clawson and Knetsch, 1966).

Given these measures of costs and benefits, proposed water resource projects with widely different purposes and in geographically dispersed locations can be ranked according to their net estimated contributions to social welfare (benefits minus costs) and other benefit-cost relationships. This information can be used to allocate the water-resource investment budget to that set of projects which maximize the net value of expected benefits.

A crucial assumption for performing benefit-cost analyses of alternatives is that the benefits or outcomes can be valued by their market prices or those of similar alternatives. Yet, the objectives of many, if not most, social programs often have no market counterpart. If a program is designed to improve the environment, how do we obtain a market price for the reduction of hydrocarbons and visible particles in the air or the reduction of pollution in water? While it is true that some of the benefits of such an action would be reflected in reduced medical costs and the added value of human lives saved, as well as in decreases in pollution-related deterioration of property, it is difficult to quantify the psychic benefits of clean air and water. How do we obtain a price for the aesthetics of open space or of conserving a rare species of bird or animal? What is the market price that will help us assess the benefits of increase in the self-concept, reading level, or music appreciation of a youngster? In each of these instances it is difficult to express outcomes in terms of their market values because a market does not exist for such services.[4]

In such situations the effectiveness of a strategy is expressed in terms of its actual physical or psychological outcome rather than its monetary value. That is, the monetary measures of resource costs are related to the effectiveness of a program in producing a particular impact. When the effectiveness of programs in

achieving a particular goal (rather than their monetary values) is linked to costs, the approach is considered to be a cost-effectiveness rather than a cost-benefit analysis. For example, one might examine various alternatives for raising the literacy level of a population, for reducing hydrocarbons in the air, for reducing infant mortality, and so on. In this context, cost-effectiveness analysis enables us to examine the costs of alternative programs for achieving particular types of outcomes, but prevents us from comparing the costs directly with benefits. That is, the cost-effectiveness approach enables us to rank potential program choices according to the magnitudes of their effects relative to their costs, but we cannot ascertain whether a particular program is "worth it" in the sense that benefits exceed costs, because the latter are generally expressed in monetary units while the former are rendered in units of effectiveness for achieving a particular impact.

The cost-effectiveness method was developed primarily by military analysts in their evaluations of weapon systems. Because it is difficult to construct benefit-cost calculations for the horrors of war, the application of the cost-benefit framework to national defense was aimed at achieving particular "objectives" at minimal cost. Such objectives might include the destruction of enemy targets, effectiveness being assessed according to the portion of the specific target likely to be destroyed by various combinations of strategies such as manned bombers versus missiles (Hitch and McKean, 1960). It is probably safe to assert that its application to other social endeavors has not been extensive in part because social evaluators have been less conscious of the importance of costs in decision making than has the Pentagon.[5]

Despite the relatively underdeveloped status of cost-effectiveness analysis as an evaluation tool, it would seem to be a more appropriate approach to many types of evaluations than cost-benefit analysis. In particular, cost-effectiveness comparisons require only that the impacts of alternative strategies along with their respective costs be derived, while the cost-benefit framework requires that we put a monetary value on the impact. Because social experimentation enables us to obtain information on the impacts of alternative treatments or programs, data on effectiveness are easier to provide. That is, the results of policy-oriented experiments or quasi-experiments lend themselves naturally to cost-effectiveness comparisons. To the degree that the effects can also be translated later into monetary values, a cost-benefit framework can be applied at a second stage. Thus, the use of the cost-effectiveness approach does permit one to do a cost-benefit analysis as well, whenever the physical or psychological outcomes can be converted into monetary measures.

For example, a study of alternative population control strategies might evaluate contraceptive techniques and educational programs by their effect on birthrates. The measure of effectiveness for each option would be the reduction in birthrates, as compared with no change or an increase in birthrates among a similar population not participating in the experiment. In this hypothetical experiment, it would be possible to arrive at the levels of effectiveness in reducing birthrates of each combination and type of contraceptive technique and educational approach, and these levels could be compared with the relevant costs. From this information the potential choices could be ranked according to their cost-effectiveness ratios.

While this exercise would enable policy makers to choose among different approaches to reducing population growth, it would not permit them to compare the productivity of using resources for this purpose with that of applying them to improvements in health, education, transportation, or nutrition. Such a comparison would require that the benefits and costs of investing in any particular area be compared with those of other areas. To the degree that reductions in birthrates can be assessed in monetary terms, it is possible to convert cost-effectiveness data into cost-benefit information. Even if the ultimate cost-benefit calculations cannot be made because of inabilities to set a value on the benefits of a program, the cost-effectiveness rankings still represent a valuable basis for choosing among programs that have the same objective. In the presentation that follows, then, the emphasis will be on cost-effectiveness methodology, although it will be stressed that many of the issues are common to both cost-benefit and cost-effectiveness approaches.

Before proceeding to examine more extensively the meaning, methodology, and usefulness of cost-effectiveness analysis, it is useful to define one more related branch of evaluation, cost-utility analysis. While cost-benefit analysis enables a direct comparison of costs and benefits stated in monetary terms and cost-effectiveness analysis represents an attempt to evaluate directly the costs of alternative ways of achieving particular outcomes, cost-utility analysis incorporates the decision maker's subjective views in valuing the outcomes of alternative strategies (Fisher, 1964: 33-48).

In concept, cost-utility analysis relies more heavily upon qualitative factors and subjective judgments than do cost-benefit and cost-effectiveness analysis. As such, it is particularly useful where a complex set of outcomes is associated with each strategy, and they cannot be assessed by market values or other measures common to all of them (Lifson, 1968). Briefly, the decision maker assesses the results of alternative courses of action according to their perceived values of utilities to him. These utilities are then expressed with respect to the costs and the probabilities of obtaining the expected results. The decision maker sets a criterion for making choices, the most common of these being the maximization of expected utility subject to a budget constraint.

II. THE COST-EFFECTIVENESS TECHNIQUE

The use of cost-effectiveness analysis in evaluation can best be shown by constructing a simple illustration of its use. From this hypothetical example it is possible to grasp the added dimension that the approach provides for social decision making. Following this illustration we will attempt to describe the principles and procedures for considering and measuring costs and effects.

Assume that we are engaged in evaluating programs designed to reduce the rate of recidivism of convicts who are released from the state prison. For purposes of this exercise let us define the rate recidivism as that proportion of former prisoners who are arrested and convicted of criminal acts within five years of being released. The existing program is one that keeps a record of the addresses and

employment circumstan-
released on parole to r
bookkeeping relationshi
or psychological counseli

To reduce an appare
wishes to consider a n
include: (1) a job placen
a program that combines
assistance of evaluation e
the impacts of the three
months all of the release
the state are assigned ra
psychological services, (3
or (4) normal existing arra

After five and one-half
five-year rates of recidivis
for those in the psycholo
program; and 37% for the
upon this appraisal it ap
existing approach, but the
by the job placement pro
circumstances the evaluati
tion that the combination
well.

Table 5.1 shows a hy
recidivism programs for
discussed above are compa
city we have assumed that
to each group. The five-y
parentheses are noted the
we can calculate the num
arrested and convicted of
prison.

Hypothetical total cost
assumed that the normal
"auditing" or "accounting
and activities of each ex-
programs are more costly,
assumed that some aspec
separate ones, so that the
the combined total of the
among the number of subj
is the least expensive and
subject.

But the average cost pe

employment circumstances of former prisoners, as well as requiring those who are released on parole to report periodically to their parole officers. Beyond these bookkeeping relationships there are no systematic attempts to provide either jobs or psychological counseling and assistance.

To reduce an apparently high rate of recidivism, the State Prison Authority wishes to consider a number of alternative programs for ex-prisoners. These include: (1) a job placement program, (2) a psychological services program, and (3) a program that combines both job placement and psychological services. With the assistance of evaluation experts, a major social experiment is initiated to determine the impacts of the three alternatives on the rate of recidivism. For a period of six months all of the released male prisoners who are returning to the major cities of the state are assigned randomly to one of four groups: (1) job placement, (2) psychological services, (3) combination of job placement and psychological services, or (4) normal existing arrangements for ex-prisoners.

After five and one-half years the experiment ends with the following results. The five-year rates of recidivism were 15% for those in the job placement program; 26% for those in the psychological services program; 12% for those in the combination program; and 37% for those ex-prisoners who received no special treatment. Based upon this appraisal it appears that all programs were more successful than the existing approach, but the combination treatment showed the best results, followed by the job placement program and the psychological services one. Under normal circumstances the evaluation might have ended here with the policy recommendation that the combination program be selected, but we wish to review the costs as well.

Table 5.1 shows a hypothetical cost-effectiveness comparison of the anti-recidivism programs for released prisoners. The three experimental treatments discussed above are compared with the results of the normal program. For simplicity we have assumed that exactly equal numbers—10,000 subjects—were assigned to each group. The five-year recidivism rates are shown, and beneath them in parentheses are noted the rankings of the results (1 is best). Based upon these rates we can calculate the number of persons who were not recidivous, who were not arrested and convicted of criminal acts within five years of their release from state prison.

Hypothetical total costs for each program are shown on the next line. It is assumed that the normal program is least expensive because it is essentially an "auditing" or "accounting" approach for maintaining information on the location and activities of each ex-prisoner. The job placement and psychological services programs are more costly, and the combination of them is the most expensive. It is assumed that some aspects of the combination program are duplicated in the separate ones, so that the cost of the combination approach is somewhat less than the combined total of the two separate components. When these costs are divided among the number of subjects in each group, it is clear that the "normal program" is the least expensive and the combination shows the highest average cost per subject.

But the average cost per subject tells us nothing about the cost for obtaining the

TABLE 5.1 COST-EFFECTIVENESS COMPARISON OF ANTI-RECIDIVISM PROGRAMS FOR RELEASED PRISONERS

	Treatment			
	Job Placement	Psychological Services	Combination of these	Normal Program
Experimental population	10,000	10,000	10,000	10,000
Five-year rate of recidivism	.15 (2)	.26 (3)	.12 (1)	.37 (4)
Number of persons not recidivous	8,500 (2)	7.400 (3)	8,800 (1)	6,300 (4)
Total cost	$10,000,000	$ 9,000,000	$16,000,000	$ 5,000,000
Average cost per subject	$1,000 (3)	$ 900 (2)	$1,600 (4)	$ 500 (1)
Average cost per nonrecidivous subject	$1,176 (2)	$1,216 (3)	$1,818 (4)	$ 794 (1)
Number of persons not recidivous in comparison with normal program	2,200 (2)	1,100 (3)	2,500 (1)	---
Additional cost beyond normal program	$ 5,000,000 (2)	$ 4,000,000 (1)	$11,000,000 (3)	---
Marginal cost per additional non-recidivous subject	$2,273 (1)	$3,636 (2)	$4,400 (3)	---

desired criterion, namely the reduction in recidivous subjects. The next line compares the average cost per nonrecidivous subject. According to this comparison, the "normal" program shows the lowest average cost per nonrecidivous subject, followed by the job placement program and the psychological services one, while the combination program seems to have the highest cost. But this comparison is not completely valid because the programs are being credited for subjects who probably would not have been recidivous even in the absence of the programs. For example, the "normal program" is merely a bookkeeping effort. While maintaining regular contact and information on the activities and whereabouts of released prisoners will have some effect on the likelihood of their returning to crime, it is likely that even in the absense of such a program a significant portion of the men would not be recidivous. Accordingly, using the total number of nonrecidivous subjects as a basis for calculations understates the cost of each "success", because it counts any nonrecidivous person as a credit to the program. This bias is most severe for the

"normal program," but it is also evident for the other groups.

These preliminary calculations lead us to the final set of results in Table 5.1. If we assume that the "normal program" is required by law, then we do not have the opportunity to eliminate it. Assuming that each of the other treatments also provides for regular contact with the person and collection of the relevant information, the policy question is what cost (for each *additional* nonrecidivous person) is saved by one of the special treatments. In comparison with the "normal program," the job placement, psychological services, and combination programs enable an additional 2,200, 1,100, and 2,500 persons, respectively (out of 10,000 subjects), to avoid returning to crime and prison. As we noted from the recidivism rates, the combination program seems most successful, followed by job placement and psychological service programs.

But the additional costs beyond the normal program vary from treatment to treatment. While the job placement program cost an additional $5 million, the psychological services cost another $4 million, and the combination program had an added cost of $11 million in comparison with the standard approach. From these data we can calculate the marginal or additional cost for each additional nonrecidivous subject. This varies from $2,273 for the job placement program to $4,400 for the combination approach. In other words, it cost about half as much to reduce recidivism by one person via job placement as via the combination program. The psychological services program was about midway in cost per additional nonrecidivous subject.

In summary, although the experiment demonstrated that the combination program showed the most success in reducing recidivism, its higher cost would not be justified. Rather, job placement would appear to be the most promising approach from the cost-effectiveness vantage point. To illustrate the impact of choice of approach on the budget of the social agency that is administering the program, we can calculate that the cost of "saving" an additional 1,000 released prisoners from returning to prison is about $2.3 million under the job placement program, $3.6 million under the psychological services approach, and $4.4 million under the combination program.

This illustration brings out a number of points. First, often what appears to be the most "effective" program may not be the most *cost-effective.* In this instance the governmental cost for each additional nonrecidivous man would have been twice as high had the most "effective" treatment been used as a basis for program selection. Second, not only may effectiveness rankings differ from cost-effectiveness rankings, but the total social costs of making the wrong choice by not considering program costs may be substantial. Third, different measures of cost may provide different implications, as a comparison of the figures on average cost per subject, average cost per nonrecidivous subject, and marginal cost per additional nonrecidivous subject show. Therefore, it is imperative that the appropriate cost comparison be used in order to obtain appropriate results.

Finally, this particular example lends itself to a subsequent cost-benefit evaluation, since it is possible to assess in monetary terms many of the benefits of attenuating recidivism. The decrease in recidivism is equivalent to a concomitant reduction in crimes against property and people. These can be evaluated, in turn,

according to the reduction in social costs associated with these crimes as reflected in decreases in property damage and the medical costs and lost earnings deriving from personal injury and death. Moreover, the increases in earnings and taxes from the employment of each additional nonrecidivous person can be estimated. Finally, the reduction in costs of the legal, penal, and police systems, as well as public assistance for dependents of the prisoner, associated with reduced crime can be calculated. All of these benefits taken together can be compared with the costs of reducing the recidivism rate to determine the value of such programs in comparison with other potential social investments.

While this illustration was used to demonstrate the usefulness of integrating the impacts of various alternatives with their costs, it did not focus on the derivation of cost data and their use, nor did it focus on the nature of the concept of effectiveness. In the next two sections we will address the measurement and application of costs, as well as some issues on measuring effects.

III. ASSESSING THE COSTS OF ALTERNATIVES

Before proceeding to a discussion of how to measure the costs of alternatives, it is important to set out the conceptual nature of costs. When one normally thinks of costs, he tends to focus on direct expenditures or what are often called accounting costs. That is, the costs of a particular action are commonly viewed as equivalent to the financial outlay associated with that activity. The concept of costs we will use is considerably broader. We define costs as representing that set of social sacrifices associated with any particular choice among social-policy alternatives.

The sacrifice or "opportunity cost" concept is based upon the economic notion of alternative uses of resources. When resources are used for one purpose, they cannot be used for other ends; therefore, the costs to society of choosing an alternative are the sacrifices implied by the "paths not taken." To the degree that these can be measured in monetary terms, we can calculate the costs associated with any alternative. In this sense the term *costs* refers to the monetary value of all the resources associated with any particular action, and their value is determined by their worth in the most productive alternative applications. Thus, the explicit expenditures associated with any particular course of action represent only a partial measure of the total costs.

A few examples of cost assessment are useful for delineating this principle. Using the criterion of costs as a sacrifice of other opportunities (or the value of foregone alternatives), we can examine the cost of a college education. It is obvious that a part of the cost of education is represented by direct expenditure on personnel, materials, and facilities. In addition, more subtle costs are imposed by the fact that individuals who enroll in colleges and universities are foregoing productive employment in the labor force. Not only does this represent a cost to them as individuals in the form of "lost" earnings, but it also represents a cost burden to society in terms of foregone production and tax contributions. Thus, the total cost of a college education includes not only the direct expenditures for education that normally come to mind, but also the implicit costs reflected in the earnings and

production foregone by persons enrolling in colleges rather than working in the labor force (Schultz, 1960).

A second example is that of programs that use contributed imputs. For example, some educational and health endeavors draw heavily upon unpaid volunteers who contribute their time and energies. While such inputs will not be reflected in the budgeted costs of the programs, they represent resource costs nevertheless since they could be used for alternative endeavors. Accordingly, we normally impute a value to these contributed inputs in order to be fully aware of their hidden worth if the budget were to reflect the full burden of resource costs.

A final illustration of the divergence between accounting costs and true social costs is the value of client time. Many social endeavors require that the client take a considerable portion of his time to obtain services. The most poignant example probably appears in the health services. Some organizations require greater expenditure of patient time, in queuing and waiting for assistance, than do others. There is clearly a cost to the client or patient in terms of foregone work and other activities. For women the cost may be reflected in monetary outlays for child care, and for both men and women in the workforce there is a potential cost in the form of lost earnings. Even when there are no direct monetary costs due to waiting, there is a sacrifice of other alternatives that would have been undertaken during that time. To the degree that the evaluation of health service alternatives does not account for these costs, it will not reflect accurately the costs of the services.

As we will see below, a group of decision makers may rationally take into account only its own costs when making a choice among alternatives, but this does not mean that costs to other constituencies should be ignored in the overall evaluation. To the contrary, all costs should be reviewed for purposes of uncovering the true social sacrifice of resources associated with a given program and level of effectiveness. Beyond this an assessment can be made of who pays the costs. While the latter is of crucial importance for the decision maker considering a particular program, the former is necessary for overall social evaluations of cost-effectiveness comparisons.

There are two basic steps in determining the costs of particular alternatives. At the first stage it is necessary to determine the specific resources required and to place a monetary value on their use. The second step entails determining which of these costs must be taken into account by any particular decision maker and how he should construct the relevant cost comparison for the application he is concerned with. These phases will be addressed in turn.

Measuring Costs

The measurement of costs would be a relatively simple task if it were only a matter of scrutinizing accounting statements and selecting the appropriate numbers. Almost never is this the case. To the degree that the cost-effectiveness evaluations are based on social experiments, it is not likely that the collection of cost data was built-in to the evaluation design. When attempts have been made to collect such information, the data tend to be crude and incomplete in comparison with data collected on other aspects of the experiment. Moreover, even if rather precise cost

data were derived for an experiment, it is not obvious that the same information could be used as a basis for a full-scale policy implementation. It is often difficult to separate the developmental costs of an experiment from the operating costs of a program; administrative costs of ongoing programs are likely to differ from their experimental counterparts; and the idiosyncracies of the experimental phase may not be duplicated in the routine of application.

In the case that cost-effectiveness comparisons will be based upon evaluations of actual programs rather than experiments, cost information may be easier to derive. Yet, as we noted above, the cost estimates may be incomplete because they are based upon expenditures rather than upon the cost of all resources that enter the program; and the classification of costs may be inappropriate for use in choosing among policy alternatives.

In the long run the best solution to these problems is to construct cost information systems as part of the social experiments and as an integral part of the social programs themselves. This procedure will enable the measurement and monitoring of costs on a continuous and relatively precise basis. It is hoped that an increasing amount of energy will be devoted to integrating cost information into experimental and program designs. Yet, until such information is collected routinely, cost-effectiveness comparisons will have to be made on the basis of ad hoc assessments. Fortunately, the conceptual basis for constructing a picture of costs is similar whether the effort is built into the program being evaluated or is designed as a separate endeavor. Accordingly, one overall approach can be used as a model for cost measurement.

The method of cost estimation advocated here will be denoted the "ingredients approach." The name derives from the focus on listing at the first step the ingredients or inputs required by the program, and assigning costs to them only after all of the ingredients are accounted for. While this priority may seem trivial at first, its importance stems from the attempt at an exhaustive listing of all of the resources utilized by a program. When one focuses immediately on costs, there is a tendency to omit from consideration those inputs that are not obvious or that do not enter budgets explicitly. The ingredients approach requires that the requisite inputs of all types required by a program be recognized for purposes of estimating program costs.

Accordingly, the first order of business is to write a description of the program and its components. From this summary one can attempt to list all of the ingredients that enter the program. Table 5.2 represents a hypothetical worksheet for estimating costs. In the right-hand column all of the program ingredients are listed by major category. These components include personnel, facilities or physical space, materials and equipment, other inputs, and the value of client time. The purpose of the ingredients column is to include exhaustively all of the elements required by the program in as detailed a breakdown as is necessary. Within each category are subcategories. Thus personnel can be divided into administrative personnel and then subdivided again into such specific titles as director, assistant director, secretary, clerical, and so on.

A separate category is set out for the value of client time and other client inputs.

TABLE 5.2 HYPOTHETICAL WORKSHEET FOR ESTIMATING COSTS

	Total Cost	Cost to Sponsor	Cost to Other Government Levels or Agencies	Contributed Private Inputs	Imposed Private Costs
Personnel					
_____					
_____					
_____					
Facilities					
_____					
_____					
_____					
Material and equipment					
_____					
_____					
_____					
Others (specify)					
_____					
_____					
_____					
Value of client time and other client inputs					
_____					
Total					
User charges		-( )			+( )
Other cash subsidies		-( )	+( )	+( )	
Net total					

As we asserted above, this resource is often neglected in cost calculations. Such an omission is equivalent to assuming that client time has a value of zero, and under such a presumption policy choices will be insensitive to the burden on the population receiving the services. One should list the number and types of man-days lost by the clients. For example, for an educational project one could provide the number of man-years of work foregone by a particular group of students. Health service programs might show the number of man-days of patient time associated with obtaining treatment, including travel and waiting time. In addition other client inputs should be accounted for in this section, for example, the transportation requirements to obtain services or the required purchase of some complementary ingredient. For example, educational programs often require the private purchase of books and materials, and many social projects have a transportation component.

Cost evaluation must consider these inputs in the total endeavor, whether the costs are borne privately or socially.

Given the list of ingredients, the next step is to estimate their costs. In the present case we will assume a one-year program. Multi-year program comparisons require certain analytical modifications that we will address in a later section. The cost of each ingredient is established either on the basis of cost experiences for that input or some other guideline.[6]

A major rule of cost estimation is to devote attention to any particular category according to the proportion of the total budget reflected by that component. Thus, if a category represents 50% of the total budget, the analyst should allocate around half of his energies to developing cost estimates for that classification. A category that represents only 5% of the total deserves no more than 5% of his attention. The reason for this principle is fairly obvious. Small percentage errors in estimating large categories can amount to large aggregate errors in cost measurement, while even large percentage errors in estimating trivial categories will have little effect on the total cost figure.

Because the personnel category typically accounts for 70 to 80% of the total budget, this component deserves the major share of attention in estimating costs. Fortunately, personnel costs can usually be appraised in a straightforward way from direct cost experience or from programs that use similar personnel capabilities. It is important to include not only salaries and wages, but also fringe benefits and other costs that employers must pay, such as contributions to health, life, and disability insurance plans, to pension plans, and to such payroll taxes as social security. These costs can be as high as one-third of salaries. The value of voluntary personnel is determined by what it would cost if such persons were paid for their services.

The costs of facilities is usually more difficult to estimate. In the simplest case, annual rental payments represent the cost. Where buildings are purchased or where the facilities used are part of a larger endeavor, the assessment of costs is more complicated. The "annual cost" of facilities that have been purchased is obtained by estimating the portion of the facilities used up (depreciated) in an annual period, as well as the opportunity cost of the investment in the sense that the investment could be obtaining social benefits in other uses. The "opportunity cost" aspect can be captured by applying an appropriate rate of interest to the net value of investment. The suitable interest rate depends upon a number of complicated factors, and the issues underlying the choice of rate have been hotly debated (Marglin, 1963; Baumol, 1968). Typically an interest rate of 5 to 10% is imputed.

The cost of depreciation is based conceptually on the fact that facilities have a limited life, so that in principle a portion of them are consumed by each year's use.[7] Thus, if a building has a 30-year life and lacks even salvage value at the end of that period, about one-thirtieth of that building is used up each year (assuming a constant annual depreciation.) In this instance the cost of depreciation would be calculated by taking one-thirtieth of the original cost of the building for each annual period. Added to this would be the opportunity cost of the investment which would be obtained by multiplying the chosen interest rate times the value of the building after present depreciation is accounted for.

To provide a simple example, assume that a project purchases a $900,000

building that has a 30-year life. Using the preceding methods of calculating the annual cost of that building, depreciation would be estimated at about $30,000 a year; and the opportunity cost of the investment would depend upon how much of the investment was depreciated. For example, if the building were 15 years old, we would calculate the opportunity cost for the sixteenth year on only $450,000, or half of the original cost of the building. At an interest rate of 5% that would amount to $22,500. (In the first year the estimated opportunity cost would be about $45,000; in the second year, $43,500; in the third year, $42,000; and so on.) In the sixteenth year, then, the cost of the facility would be $30,000 for depreciation plus $22,500 for the opportunity cost of the investment, with a total annual cost for that year of $52,500.

The measurement of facility costs may require fairly extensive computations if facilities are shared with other programs or are provided by a parent enterprise. In the former case the facilities and other resources may be shared by a number of activities of which only a particular one is the focus of the cost-effectiveness study. In the latter case the facilities may be inseparable in an accounting sense from the larger entity. An illustration of this phenomenon is the determination of the value of high-school facilities devoted to a school dropout program, or the value of the facilities used for an outpatient clinic located in a hospital.

There are several methods of estimating the costs of such facilities. If the total cost of shared facilities can be ascertained from the total rent or by estimating the depreciation and opportunity cost of investment of those that are owned, it is usually possible to prorate these costs among the individual programs. The basis for prorating costs should be some common denominator for all of the programs such as the proportion of the total space used by each program or some other appropriate guideline. This approach can also be used when facilities are provided by a parent enterprise such as a hospital, school, or community center. However, in some cases it is difficult to obtain the information required to prorate costs since this procedure requires data on other projects or on a parent organization that may not be participating in or cooperating with the evaluation. Under such circumstances, it is often preferable to use the alternative method of imputing a rental value equal to the annual rent for comparable facilities. As an illustration, assume that space of a comparable quality can be obtained in a similar location for about $2 per square foot per year. If the program that is being evaluated uses about 1,000 square feet, the imputed cost of facilities would be about $2,000 per annum.

In general, materials and equipment and other inputs can be costed on the basis of expenditures or on the basis of their market values if they have been contributed. The value of client time and other client inputs is based either on direct expenditures of clients where they are evident or on an assessment of the worth of client time. The former category may include client expenditures on books, materials, and other educational inputs in the case of schooling, or the privately borne costs of health programs, such as prescriptions and transportation. The latter category is meant to assess the loss of productive time from other activities, whether in the labor market or in the household. One rule of thumb that is often used to value client time is that of appraising it at its market value with respect to earnings foregone. In order to do this it is necessary to determine the age, racial,

and sexual composition of the clientele. From these breakdowns it is possible to use Census data on earnings to determine opportunity cost (Becker, 1964.)

Treatment of Time in Cost Measurement

Where the program requires more than one year of operation to obtain effects, it is necessary to estimate the costs for each year. There is no problem if the costs do not vary from year to year; however, many multi-year programs required additional allocations in the initial period for training and other start-up costs. Other programs may entail higher costs in the later years. Merely to add annual costs for several years does not provide an appropriate basis for comparing multi-year programs whose costs have different time patterns. Programs that allocate the bulk of their costs in the earlier years are not directly comparable with those that allocate the bulk of their costs in later years because the former entail a higher opportunity cost by sacrificing resources earlier in time; thus those resources are withdrawn from alternative uses for a longer period than when they are used later in the program cycle.

As an illustration, let us consider two alternative programs that show total costs of $500,000 for a five-year period. Assume that program A spends $200,000 for the first year and $75,000 for each of the next four years. Assume that program B spends $50,000 for the first four years and $300,000 for the last year. While both programs expend $500,000 over a five-year period, the social burden is greater for program A.

One way of understanding this concept is to consider the projects as bank accounts with $500,000 in initial deposits. Since account A is drawn upon earlier than B, it will yield less interest, where the interest that is derived represents the productivity of investments of the bank. That is, when the money is spent on the program, it cannot be used for other productive endeavors which the bank would invest in. If we assume that the expenditure transactions that we noted for each year would be consummated on the first day of that year, bank account A (or program A) would accumulate about $43,500 in interest payments at a 5% rate of interest, compounded annually over the five-year period. Bank account B (or program B) would accumulate about $85,500 in interest over the same period.

In short, the expenditure pattern reflected in program A is more costly to society as reflected in foregone interest payments which in turn reflect the productivity of investment in alternative endeavors. (This is one of several places where the economic argument seems to cut corners. The money A spends is back in circulation doing its productive thing at year 1.) The earlier expenditure commitment of program A entails a greater social sacrifice of resources than that of program B even though both spend a total of $500,000 over a five-year period. To make the two comparable with respect to the time allocation, we calculate the present values of their expenditure streams by discounting future costs by a rate of interest that reflects their value in alternative endeavors. The formula for obtaining the present value of a stream of costs is:

$$PVC = \sum_{t=1}^{n} \frac{X_t}{(1+i)^t}$$

where PVC represents the present value of a stream of costs; X_t denotes the cost for year t; n represents the final year in which costs are incurred; and i represents the appropriate rate of discount or interest. As we noted above, this rate of interest is supposed to reflect the rate of return to society of alternative uses of resources, but the selection of a particular rate depends upon a number of judgments (Baumol, 1968). Most studies utilize a rate of 5 to 10%; however, it is wise to select a number of rates within the appropriate range in order to ascertain the sensitivity of the result to the choice of discount rate.

Table 5.3 applies the present value approach to assessing costs of the two programs presented above. The first and third columns of the table show the allocation of annual costs over the five-year time span for the two alternatives. The second and fourth columns convert each of these annual "future" costs to present values by applying a 5% rate of interest or discount rate to the present value of cost formulation. Thus, the cost for each future year is shown according to its present

TABLE 5.3 CALCULATIONS OF PRESENT VALUE OF COST STREAMS FOR TWO PROGRAMS

	Program A		Program B	
	(1) Annual Cost	(2) Present Value $\frac{Xt}{(1+i)^t}$	(3) Annual Cost	(4) Present Value $\frac{Xt}{(1+i)^t}$
Year 1	$200,000	$200,000	$ 50,000	$ 50,000
Year 2	75,000	71,429	50,000	46,619
Year 3	75,000	68,027	50,000	45,351
Year 4	75,000	64,789	50,000	43,193
Year 5	75,000	61,703	300,000	246,812
Total	$500,000	$465,948	$500,000	$432,975

value equivalent. Because we assumed that the costs in each year would be incurred on the first day of that year, the present value of the first-year costs is also the annual cost. Each subsequent year, however, is divided by the factor $(1 + i)^t$ where t is set to equal to 1 for the annual cost incurred in the second year, by virtue of the fact that it is allocated on the first day of the second year. When we add the separate present values in columns 2 and 4, we find that the total present value of costs for program A is about $466,000 while for program B it is only $433,000. If the programs showed equal effectiveness, program B would be a preferable policy choice because of its lower costs. In general, any comparison of programs characterized by different time patterns of costs will necessitate the conversion of the annualized costs into present values in order to make them comparable among programs.

Uncertainty and Cost Measurement

One other general aspect of cost measurement that should be considered is that of uncertainty. In some cases it may be extremely difficult to ascertain the costs of

particular program ingredients. This is especially likely to be true in the assessment of programs, constructed on the basis of piecemeal experimental results, which have never been initiated as total programs. To the degree that the uncertainty applies to only a small fraction of the total ingredients (e.g., non-personnel inputs only), the total cost picture might not be affected appreciably by errors in estimating the costs of these inputs.

The usual treatment of uncertainty is to use the most comparable experience available as a basis for cost estimation. Such comparable data can be adjusted for idiosyncracies of the present application. Where the margins of error appear to be extremely large, it is preferable to suggest a reasonable upper boundary and a lower one. Different values can be used within this range to see to what degree the total cost estimates are sensitive to different assumptions. If total costs seem highly sensitive to the assumptions on which the cost estimates for particular components or ingredients are constructed, it is best to use several alternative assessments of costs in the final cost-effectiveness analysis to see if they affect the results. Normally a high and low value are estimated, and the midpoint or some other value within this range is selected as the most likely figure.[8]

Appropriate Cost Comparisons

As we noted in the illustration on criminal recidivism, cost comparisons are possible. Each is appropriate or inappropriate depending upon the questions one wishes to ask and who is asking the question. It is especially useful to review the allocation of costs among different social entities. From the point of view of society, the total cost of each program must be considered alongside its impact in order to make choices among programs. That is, the most efficient use of the society's resources is made when those programs are selected that show the greatest effects relative to the social sacrifice of resources that they entail. Certainly, an evaluator who is concerned with the largest social perspective will wish to consider the total social costs associated with each alternative.

A public agency may be much less concerned with total social costs than with its own cost burden. That is, a particular governmental unit will be interested primarily in maximizing the impact of its own resources. Its relevant cost comparison will be among the costs it incurs for each alternative, and there will be a tendency of the sponsoring agency to ignore costs borne by other entities.

The worksheet in Table 5.2 is designed to classify the costs among several of these entities including the sponsoring agency, other levels of government and government agencies, and the private sector. For reasons that will be explained, the private sector is divided into contributions and imposed private costs, especially those borne by clients. Because each of these groups may evaluate the "cost-effectiveness" of a program on the basis of its own costs (and benefits), such a classification is rather important for understanding policy choices and political support for particular programs that may appear to be inefficient when total costs and effects are reviewed.

Higher levels of government or other government agencies may contribute to a program by providing particular program ingredients or cash subsidies.[9] For

example, some states provide the salaries for reading specialists used by local educational agencies. Another example of a contributed input is the surplus federal properties and equipment given by the federal government to other governmental agencies for public purposes. Cash subsidies are illustrated in public assistance monies used to support housing programs for low-income families. Other examples include the federal subsidies to local governments for acquisition of parks.

Private costs can be divided into contributed inputs and private cash contributions. The former include the value of time of volunteers (especially in hospitals, schools and libraries) and donated facilities, materials, and equipment. In addition, there is that large category of imposed private costs, the direct expenditures required of clients for obtaining the service. These represent a cash subsidy from the client to the sponsoring agency. Examples of user charges include the costs that clients must pay for health services, admissions fees to museums and other public institutions, and tuition and fees at public educational institutions.

To determine the net costs to these entities, both contributed inputs and cash subsidies must be taken into account. The various columns to the right of the Total Cost column enable the costs of ingredients to be allocated among the four classifications. Thus, the totals at the bottom of the columns represent the costs associated with the provisions of the ingredients in the program by each group. But, in addition to these there are cash subsidies to the sponsor from both government and private sources as we described above. Accordingly, at the bottom of the appropriate columns we must add or deduct the value of these cash subsidies to determine the net costs to each constituency.

This adjustment is implemented by adding user charges to the Imposed Private Costs and deducting this subsidy from the Cost to Sponsor. Likewise the case subsidies provided by other governmental levels and agencies and by private contributions should be added to the costs in those columns and should be deducted from the sponsor's costs. These transactions are symbolized by the parentheses with accompanying plus or minus signs. When these accounting transfers are made in the table, the results show the net total costs to each entity. After these adjustments it is obvious that the net cost to the sponsoring agency can be considerably less than the total cost of the program. If a program with relatively high cost-to-effectiveness ratios is more highly subsidized than one with relatively low cost-effectiveness, there is the risk that a sponsoring agency will select the program that is less appropriate from the perspective of the larger society. That is, the cost-effectiveness ratio *to the sponsoring agency* for any program is improved by a larger subsidy. The importance of this point is to illustrate the cause of possible divergence between agency optimization and the maximization of social objectives. Cost-effectiveness evaluation can play an important role in documenting the contradiction.

Other levels of government and private constituencies would opt rationally for programs that minimize the costs to their entity for any given level of program benefits. Thus, it is not difficult to see that each group, comparing its costs to the perceived effectiveness of particular programs, may calculate different priorities. These kinds of conflicts might even be more apparent if we were to analyze the

distribution of the tax burden and benefits of particular programs (Gillespie, 1965). The former is generally beyond the scope of a cost-effectiveness analysis, and the latter will be reviewed in the later section on effectiveness.

Cost-effectiveness comparisons are made on the basis of cost compilations that might analyze costs in any one of several ways. The most common of these are: total costs for obtaining a given level of effectiveness, average costs per unit of effectiveness, and marginal costs for additional units of effectiveness.

The comparison of *total program costs* is most appropriate when we are confronted with alternatives of about equal effectiveness. For example, assume that we are confronted with three possible routes for a highway planned between two remote points. Each route has about the same distance, but they differ with respect to the types of construction and the types of property that will have to be condemned enroute. In that case the evaluator might simply wish to calculate the total cost of each alternative and select that choice which minimizes the total project costs.

When programs differ in terms of their effectiveness, it is usually advisable to compare the *average cost* per unit of effectiveness. For example, assume that we evaluate several approaches to decreasing infant mortality, including prenatal educational programs, different approaches to maternal prenatal services, and nutritional supplements. Every woman is assigned randomly to a treatment or to a control group, and the levels of infant mortality are calculated for the groups. The effectiveness of a program can then be stated in terms of the reductions per standard unit (e.g., 1,000 pregnancies) of aborted births and of infant mortality through the first year of life. The programs may show widely varying results and widely varying costs. Accordingly, dividing the reduction in infant mortalities associated with each treatment by the cost of the treatment yields the cost per standard reduction measure.

This average cost per unit of effectiveness has the very desirable quality of permitting cost-effectiveness comparisons among divergent programs with very different characteristics as long as the total costs of the programs are available and the outcomes of the programs are measured in the same effectiveness units. For example, using the concept of average cost per unit of effectiveness, it is possible to compare different remedial-reading programs in terms of the cost per additional point of reading score. Delinquency prevention programs can be evaluated according to their cost for reducing delinquency rates. Health programs can be contrasted on the basis of the cost per standardized reduction in the incidence of diseases.

While the average cost measure is extremely attractive because of its generalizability across widely divergent programs, it is characterized by one major problem. Such a criterion implies that the most efficient ranking of programs at one level of the output scale pertains also to other levels of output. In fact, often programs that show low average cost at modest levels of output will show relative high average costs at greater levels of output and vice-versa. A reason is that technologies vary according to their relative intensity of fixed costs. Fixed costs represent those basic ingredients that must be acquired in order to operate the program. They are invariant with respect to the level of output of the program to the degree that they

must be obtained simply to offer *any* level of output, low or high. Accordingly, programs with large components of fixed costs will show relatively high average costs for low levels of output and low average costs for higher outputs.

As an illustration of fixed costs, consider a clinic designed to handle 400 patients per day. The program ingredients include the minimum physical facilities and specialized personnel necessary for operation. Even if only 100 patients a day are serviced, it is difficult to reduce these basic costs; but the average cost per patient will be about four times as high as at full capacity. Variable costs differ according to the number of patients serviced. For example, the number of X-rays, injections, and other patient services will be reduced with lower patient loads, but these costs are likely to be proportionate to patient load. That is, the average variable cost per patient is not strongly affected by scale of operation. If such a clinic were contrasted with a program that simply sent patients to community practitioners, it is possible that, for small numbers of patients, the cost per treatment would be lower for the latter alternative, but, for larger numbers of patients the "economies of scale" reflected in increasing utilization of the fixed overhead of the clinic would likely reduce the average cost of treatment below that for individual practitioners.

Figure 5.1 represents these cost relations. For relatively low numbers of patients it is possible to utilize individual practitioners at a relatively modest cost, but with larger numbers of patients the costs rise gradually because of the need to draw upon the more expensive practitioners in the community and also because the higher demand for practitioner services might stimulate higher prices. In contrast, the cost per patient treatment at the clinic would tend to be very high with few patients, but the cost per patient declines as the number of patients rises. Below about 200 patient treatments per day, the individual practitioner approach shows a lower cost per patient than the clinic, but beyond that level the clinic tends to be more cost-effective.

Comparisons of average cost per unit of effectiveness should be made only at

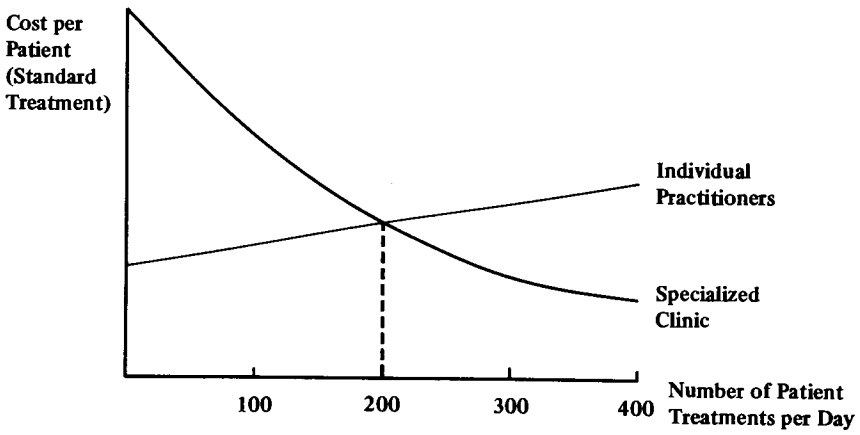

FIGURE 5.1 Hypothetical Average Cost Relationships for Individual Practitioners and a Specialized Clinic

that level of scale that is appropriate to the particular application being considered. Without taking the scale of a program into account, the results may be very misleading. For example, educational television and radio may be much less costly alternatives for teaching reading than their more traditional instructional counterparts for state and national educational systems with several hundred thousand students. But their high fixed costs make educational television and radio rather costly for teaching reading to student populations of a few thousand.

In some instances it is the additional cost or *marginal cost* per effectiveness unit that is the relevant measure of comparison. This is particularly true where average cost per unit of effectiveness changes according to the scale of program. Assume that the decision maker is faced with the choice of expanding existing programs or initiating another one. In such an event he may be less concerned about the average cost per unit of effectiveness than he is about the additional cost per effectiveness unit entailed in enlarging the existing alternatives or in creating a new program. The measurement of costs would address only those additional ingredients needed to expand the programs from their present levels, and these would be contrasted with the additional units of effectiveness that would result. This marginal-cost approach considers not all the program costs and effects, but only those at the margin. The same criterion can be used for program contraction where the decision maker wishes to choose that combination of programs that will ensure the largest reduction in the budget for a unit decrease in output.

Marginal cost-effectiveness comparisons are also appropriate when decision makers must allocate budget increases or decreases among competing programs (Schultz, 1969; Wildavsky, 1964; Rivlin, 1971). To maximize the impact of an increase, they will wish to allot additional funds to those programs that show the largest gains in output per unit of cost. Conversely, decreases in the budget should be spread among those programs that show the smallest loss of output for a given reduction in support. In these circumstances the analyst must review all potential programs according to how marginal increases or decreases in costs will affect the magnitudes of increases or decreases in output. In this analysis one need not consider all costs and effects, but only marginal costs and effects.

IV. MEASURING EFFECTIVENESS

Because most evaluation research tends to ignore cost analysis, we have devoted a substantial portion of this effort to the explanation of methods of measuring and analyzing costs. The existing approaches to social evaluation do have much to say about concepts and measures of effectiveness. Before evaluation studies proceed, they must select and operationalize a criterion. This criterion is the outcome or measure of effectiveness that is observed for all the alternative programs or strategies. We will not address the conceptualization and measurement of the criterion to the degree that we have reviewed the cost aspects. Nevertheless, three rather general issues are important to consider: (1) the time pattern of results, (2) measuring multiple outcomes, and (3) distributing effects among different populations.

The Time Pattern of Results

Just as the time allocation of costs may differ among programs, so may the time pattern of results. Consider two reading programs designed to bring "slow-readers" up to grade level in reading achievement. Program A emphasizes computer-assisted instruction and other technological aids, while Program B concentrates on small-group instruction and tutoring. Students are assigned randomly, and they remain with the program until they reach their grade equivalent in reading, at which time they are transferred to regular reading programs. At the end of five years it is observed that 80% of the children who were enrolled in program A reached this criterion compared with 70% of the children in program B. It is found that similar students who do not receive the special treatments do not obtain grade level reading proficiencies over the same period.

One possible way to carry out a cost-effectiveness comparison is to compare the total costs for the two programs with the total number of children brought up to grade level in reading within five years. This would yield the average cost—properly discounted to obtain its present value—for each program "success." But on further scrutiny we observe that the timing of "successes" varies as shown in Table 5.4. According to this presentation, program B shows relatively earlier successes than program A even though the results for program A surpass those of program B at the end of five years.

The earlier a student is able to read at grade level, the sooner he is able to benefit from regular instruction and to participate in standard reading activities for his age. If society places a premium on bringing students up to that level of proficiency quickly, it is clear that a comparison of the two programs should take the time pattern of impact into account. This adjustment can be made by weighing more heavily in calculations of effects the earlier successes than the later ones.

In the selection on measuring costs we presented a technique for doing this, the application of a discount rate to the stream of annual costs. The same approach can be used for making different time streams of effects more nearly comparable. That is, the additional number of children that are able to achieve at the appropriate reading level can be calculated for each of the five years, and the "present value" of this five-year pattern can be estimated by utilizing an appropriate discount rate much as the present value of costs was determined. The cost-effectiveness comparison of the two programs can then be based on present values of both costs and effects. In this example, the application of "present value" methodology would

TABLE 5.4 PROPORTION OF STUDENT COHORT WHOSE READING ACHIEVEMENT HAS REACHED THE MEAN ACHIEVEMENT AT THEIR GRADE LEVEL

	Program	
By End of Year	A	B
1	10	30
2	25	40
3	40	50
4	70	60
5	80	70

improve the relative standing of program B in comparison with using an approach that ignores the time pattern over which program outputs occur. In any endeavor where deferred outcomes are less preferable than present ones, it is wise to ensure that cost-effectiveness comparisons are based upon the appropriate adjustments.

Effectiveness Measures for Multiple Outcomes

Virtually all of the examples we have used envision the effectiveness of a program as a single output reflected by a single measure, but many social programs are characterized by broad aims that can be reflected in a large number of measurable (and unmeasurable) outcomes.[10] Educational programs that are designed to improve a child's achievement in one subject may also influence his performance in other subjects, as well as his attitudes (Levin, 1970b.) Delinquency prevention programs may not only decrease the amount of juvenile crime, but they may also increase the educational attainments, employability, and welfare of the young adults involved.

The focus on a single criterion measure of effectiveness ignores the other effects of programs. If the other impacts have no value, then there is no harm in omitting them. In all too many cases, however, the other effects are of some importance, and should be considered explicitly. When one considers more than one output, he is faced with the problem of how to integrate them in the cost-effectiveness framework.

Assume, for example, that we are asked to evaluate the effects of two instructional programs on *both* mathematics and reading scores. Table 5.5 shows two sets of hypothetical outcomes. In Case 1, Program A shows greater effectiveness per unit of cost than Program B for both reading and mathematics achievement. Accordingly, the evaluator would recommend the adoption of Program A. But in Case 2, Program A shows a superior effect relative to cost for reading, while Program B is more effective for mathematics. Which program is preferable from the cost-effectiveness vantage? It is obvious that there is no answer a priori. If Program A is chosen we will obtain higher reading scores, but will sacrifice 2 points of mathematics achievement. If Program B is chosen, mathematics scores will be higher, but reading scores will be less than optimal.

TABLE 5.5 HYPOTHETICAL TEST SCORE GAINS PER STUDENT FOR A GIVEN BUDGET INCREMENT TO EACH OF TWO PROGRAMS

	Case 1		Case 2	
	Reading	Mathematics	Reading	Mathematics
Program A	2	3	2	1
Program B	1	2	1	3

The only way we can obtain an unambiguous cost-effectiveness ranking between the two programs is to express the value of the outputs in common units. One of the advantages of cost-benefit analysis is that outputs are converted into monetary values. But it is often difficult to translate changes in attitudes, test scores, and

improved health into monetary benefits. The alternative is to assign arbitrary weights to the outputs so that they can be aggregated into an effectiveness index. Because this requires an ethical or normative judgement that may not express the view of everyone concerned, it is best to use several alternative weighting schemes. This approach enables each person to use his own values in scrutinizing the results, and it permits the evaluator to see whether cost-effectiveness rankings of the programs change as output priorities change (Azzi and Cox, 1973; Weisbrod, 1968).

Table 5.6 shows hypothetical effectiveness ratings for the two educational programs represented in Case 2 of Table 5.5. The test score gains for a given budget increment are translated into an effectiveness index by applying weights. When reading scores are assigned a weight of 1 and mathematics scores are assigned a weight of 2, the composite effectiveness score for Program A is 4. The comparable weightings for Program B provide a composite score of 7. Given those weights, Program B is the more effective. If reading gains were valued three times as much as mathematics gains, Program A would show a value of 7 in contrast with a value of 6 for Program B.

To compare program effectiveness the higher rating for any set of weights is underlined and equal ratings are encircled. The display can be summarized by saying that only when the decision maker believes that reading gains are at least three times as important as mathematics gains will he choose Program A. Under other conditions he will either be indifferent to the two programs or choose B. This

TABLE 5.6 EFFECTIVENESS RATINGS OF TWO EDUCATIONAL PROGRAMS UNDER ALTERNATIVE WEIGHTING SCHEMES

Case 2

Program A
Reading Weights

Mathematics Weights	1	2	3	4
1	3	(5)	7	9
2	4	6	8	(10)
3	5	7	9	11
4	6	8	10	12

Program B
Reading Weights

Mathematics Weights	1	2	3	4
1	4	(5)	6	7
2	7	8	9	(10)
3	10	11	12	13
4	13	14	15	16

approach forces the user to make an explicit ethical decision about the values of the outcomes and enables him to see the cost-effectiveness implications of his value choice. It also permits him to see how sensitive the rankings of the programs are to deviations from his weighting. The usual alternative is to ignore additional outcomes other than the particular one under scrutiny. That method places an implicit value of zero on other effects, and such a value judgment is (by default) not ordinarily obvious to the user of the evaluation findings. Accordingly, the explicit weighting scheme would seem to be a far superior approach for the multiple outcome program comparison.

Distribution of Effects Among Different Populations

A final aspect of the assessment of effectiveness is to evaluate who receives the benefits of the program. The use of a single effectiveness index for each outcome (e.g., the mean) assumes that the comparable programs have the same distribution of results across the population or that the distributions do not really matter (Azzi and Cox, 1973). The distribution of effects should be evaluated if a movement toward equality of outcomes is desirable. An example is found in education where there is a concern not only with the improvement in achievement test scores associated with a given approach, but also with the distribution of the scores (Block, 1971). In general, a program is considered to be more desirable the greater its ability to produce achievement gains and to reduce the variance in student achievement.

Assume a comparison of two instructional programs with equal average gains in student achievement. On closer scrutiny the first program has increased the test scores of all students by about the same amount, while the second has increased substantially the test scores of the top third of the student population with only modest or insignificant changes among the lower two-thirds. If we were to ignore the distribution of the results, we would rate the two programs equally effective. By attaching any positive value to the distributional consequences of the programs, we would prefer the first program over the second one for the same cost outlay.

Accordingly, when programs with equalization goals are evaluated, some attempt should be made to examine distributional changes as well as the average effect. This can be done in a number of ways (Jamison, et al., 1971). Perhaps the simplest way to measure distributional effect is to compare the change in the variance of the outcome, as well as the change in its level. Other distributional measures may be more appropriate depending upon the desired properties of the indicator (Atkinson, 1970). Given both a measure of change in outcome and its distribution, it is possible to obtain overall indices of effectiveness of programs by weighting the two according to decision priorities. This can be done by treating the measures as multiple outcomes and applying value weights to them in the manner suggested in the previous section. That is, the relative value of changes in the distribution and level of outcome can be assessed by setting out an analysis similar to that in Table 5.6.

Distributional concerns are important also when the production of benefits for one segment of the population is considered to have greater value than equal

benefits for other segments. While this judgment is inherently an ethical one (as are all distributional considerations), it can be said that many programs express special concern for poor, aged, and minority populations. For example, consider housing rehabilitation programs that make use of low-cost loans and other subsidies. The effectiveness of such programs might be stated in terms of the number of housing units rehabilitated by such public endeavors. Then the approach that provides the greatest number of rehabilitations for a given cost would appear to be the most cost-effective. But what if the groups who are the primary beneficiaries of each program differ in composition? Under one program it is essentially middle- and upper-class persons who have purchased dilapidated buildings at highly subsidized interest rates; they refurbish the buildings to use as domiciles, or to rent and sell for profit. The other program has helped the poor, minorities, and the aged to form community corporations that systematically renovate the buildings for housing rentals to members of the communities.

If the total number of rehabilitated units were the only criterion of effectiveness, it is likely that the former program would be found to be more effective than the latter. To take the distribution of benefits into account, one can treat each distributional variable as an outcome of the program and weight the variables into the effectiveness index. In this case the proportion of direct recipients or participants who are classified as aged, minority, or poor would be ascertained, along with the number of housing units rehabilitated. These can be valued by alternative rating schemes into overall effectiveness ratings that can be compared with costs.

While the distributional aspects of social programs have been neglected in the past, to ignore them is to make distributional judgments nevertheless. Such an omission represents a tacit acceptance of the existing distribution of benefits of the programs that are being evaluated. Moreover, if the housing example is realistic and representative, a cost-effectiveness analysis that does not consider the distributional aspects of the results may be systematically biased in favor of the middle and upper classes.

V. ILLUSTRATIONS AND CONCLUSIONS

Now that the rationale and methodology for cost-effectiveness analysis of social programs have been presented, it may be useful to review studies that have utilized this analytical tool. The nuances of the technique are much better reflected in examples than in an abstract description of methods. The fact that cost-effectiveness analysis has not been widely employed for the analysis of social programs handicaps this endeavor; most of the applications have been in the area of military strategies and acquisition of weapon systems (Goldman, 1967; Quade, 1964). Cost-benefit approaches have typically been used for the evaluation of public policy in other areas, perhaps because economists prefer the wider comparability of cost-benefit results. Fortunately, many of the issues raised by cost-benefit techniques are pertinent to cost-effectiveness studies.

In recent years a number of survey articles, bibliographies, and volumes of readings have been published on the subject. The most comprehensive survey

articles are those of Prest and Turvey (1965), and Musgrave (1969.) An excellent annotated bibliography is found in Hinrichs and Taylor (1969), and most of the other sources that are cited here offer extensive bibliographies. Among collections of readings on cost-benefit analysis, the most comprehensive are Dorfman (1965), Chase (1968), and Niskanen et al. (1973). These readings cover a wide variety of topics, including recreation, agricultural price supports, transportation, manpower training programs, nutrition, education, health, and methodological issues. Goldman (1967) presents a set of essays on cost-effectiveness analysis in decision making. While some of the issues treated are general, most of the papers concern the application of cost-effectiveness analysis to military decision-making.

Health Studies

Health programs represent a fertile area for cost-effectiveness and cost-benefit inquiries. Analyses in the health area have been especially varied with respect to subject and methodology. Provocative discussions of the problems in the economic evaluation of health endeavors and suggested directions for their solution are presented in Mushkin (1962) and Weisbrod (1961). The measurement of health is discussed in Hennes (1972).

Among studies of particular health topics, Maidlow and Berman (1972) attempted a benefit-cost analysis of alternative means of treating heroin addicts. A part of their evaluation was based upon the probability that former addicts would return to addiction. Attempts to estimate the costs of particular diseases are presented in Rice (1968) and tied to a general methodology. The costs of mental illness are estimated by Fein (1958), and the costs of syphilis, by Klarman (1965). These two works consider conceptual aspects of measurement.

The economics of preventing infectious kidney diseases is the focus of a study of Menz (1971). This inquiry evaluates the costs and results of a hypothetical program that would be differentiated by several different levels of intervention. Cost-effectiveness of periodic health examination strategies is explored by Forst (1973). The delivery of health services is explored in the work of Smith, Miller, and Golladay (1972), which used a cost-effectiveness framework to determine the optimal roles of paramedical personnel in the production of primary medical care.

Education and Manpower Training

Numerous cost-effectiveness and cost-benefit studies have been undertaken in the evaluation of educational and manpower training programs. In the area of elementary and secondary education, Corazzini (1968) and Hu, Lee, and Stromsdorfer (1971) have attempted to compare the costs and benefits of secondary vocational programs with general or comprehensive programs. Each of these explorations reports the conditions under which each of the alternatives might be considered superior. Levin (1970a) carried out a cost-effectiveness analysis of teacher selection where the measure of effectiveness was the increase in verbal acheivement of sixth graders. It was found that the use of more experienced teachers was five to ten times as costly as selecting teachers who showed higher verbal scores on a vocabulary test. Perl (1973) carried out a similar type of analysis

for secondary schools. According to his results, the selection of more educated teachers would provide a considerably greater impact on student test scores per dollar of expenditure than reducing the average class size.

A study of the cost and performance of computer-assisted instruction for disadvantaged children (Jamison et al., 1971) deserves special mention. This investigation compared the test gains of children in three programs that utilized computer-assisted instruction (CAI) with those of children who received conventional instruction. Two of the programs focused on mathematics achievement, and the other concentrated on reading skills. In general, the students in the CAI programs outperformed their counterparts in the control groups. The authors also evaluated the distribution of the gains in test scores to see if gains were more (or less) variable in the CAI than in conventional instruction. The CAI students tended to show more nearly uniform gains than did their counterparts. Finally the relative costs of the CAI and conventional instruction were evaluated. The combination of experimental methodology with considerations of the distribution of results and costs makes this study an especially valuable reference.

Studies of dropouts (Weisbrod, 1965), preschool programs (Ribich, 1968: 83-97), and compensatory education expenditures for children from low-income backgrounds (Ribich, 1968: 83-97) have been evaluated with respect to the estimated increase in lifetime income associated with the effects of such programs. It has been found that in general, the costs of such programs exceed the benefits when income production is used as the criterion of success. Broader evaluations of education as a social investment have found relatively high returns in both the United States (Becker, 1964; Hanoch, 1967) and other countries (Psacharopoulos, 1973).

Manpower training programs have been evaluated by cost-benefit and cost-effectiveness techniques; an important and comprehensive review is found in Goldstein (1973). Ribich (1968: 34-50) compared a number of manpower training programs. In general, it appears that the benefits of such programs far exceed their cost. A comparison of particular approaches to "neighborhood youth corps" is presented in Somers and Stromsdorfer (1972). That evaluation included an analysis by race and sex of student of the effects of both in-school and summer youth corps programs with respect to a number of criteria including high-school graduation.

Other Studies

A wide variety of other subjects has been the focus of cost-effectiveness approaches. Douglas and Tweeten have suggested a method for calculating the cost of controlling crime. Empirical findings on the factors affecting the costs of local police services and fire services in the St. Louis metropolitan area are reviewed in Hirsch (1970: 170-171). A study of the effectiveness of correctional programs is found in Robinson and Smith (1971). Since no correctional program was found to be more effective than any other, it is likely that the least costly one would be preferred; but Robinson and Smith do not present the costs of the alternatives. Empirical analysis of the costs of different service levels of refuse collection at the municipal level is presented in Hirsch (1965).

A substantial number of studies have evaluated the provision of recreational services and facilities. A general review of the issues and methodology is found in Clawson and Knetsch (1966). Specific studies of interest include Mack and Myers (1965) and Krutilla and Cicchetti (1973). Highways and transportation have been the focus of numerous analyses. The evaluation of urban highway programs was undertaken in Mohring (1965) and Kain (1967). Navigation improvements are scrutinized in a cost-benefit framework by Haveman (1973). Cost-benefit studies of urban renewal by Rothenberg (1967) and of infant nutrition by Selowsky (1973) represent exceedingly provocative applications of these analytical tools to those areas.

Conclusion

The case for carrying out cost-effectiveness analyses of social alternatives is a strong one. Given limited resources to allocate among competing programs, evaluation for social choice must consider both the costs and the results of particular approaches. Yet this branch of inquiry, like most aspects of evaluation, is a relatively young one. The conceptualization and measurement of both costs and outcomes have not and probably cannot be routinized. Accordingly, the judgments of the evaluator in setting out decision rules and guidelines for estimating costs and effects represent a crucial variable in determining the outcome of the evaluation. The omission of particular cost components or program outcomes, the selection of a particularly high or low discount rate for future costs or results, and the method of estimation of costs of program ingredients all represent areas where different judgments may alter appreciably the cost-effectiveness ratings of alternatives (Williams, 1973). The preoccupation with means-ends relations at the expense of considering processes also represents a bias of the approach (Tribe, 1973).

The implications of the foregoing are that the cost-effectiveness analyst and the user of his results should feel obligated professionally to use the tool with wisdom and caution. In particular, the analyst is urged to make his assumptions about cost and output measurement as explicit as possible with a discussion of alternative assumptions and the likely effects that they would have on findings. Where feasible the calculations should by presented on the basis of alternative assumptions so that the reader can choose those that are consistent with his own understanding of the issue. The user should recognize the fact that the conclusions of such studies do not in themselves define a policy action. Rather they serve as useful—and, one would hope, potent—sources of information that must be combined with factors that have not been taken account of in the cost-effectiveness inquiry, in order to make public choices that are sensible, efficient, and equitable. In this context, cost-effectiveness analysis can be a powerful and productive ally.

NOTES

1. For an overview of public budgeting systems and analysis, see Lee and Johnson (1973.)

2. This estimate was calculated from data provided by Dean Jamison. The actual cost of the program that was used as the basis of the experiment was considerably higher because it was designed for research and development purposes rather than routine application.

3. It is not the purpose of this article to review the enormous theoretical and practical difficulties in deriving allocation decisions that will maximize social welfare. The classical treatment of the subject is found in Pigou (1951), and an excellent critique is Little (1957). Also see the important theoretical dilemma presented by Arrow (1951). The relationship between prices and costs on the one hand and value on the other is a rather complex one. In a capitalist society it is explained primarily by the theory of markets. A fundamental exposition can be found in Samuelson (1970), chaps. 20 and 22-25. The neoclassical treatment is found in Hicks (1946), chaps. 1-5. In recent years the realism of the theory has come under increasing criticism. See for example, Gintis (1972) and Galbraith (1970).

4. While it is possible to attempt to determine how much people would pay for such benefits, it is not clear that they could evaluate accurately the values of services that are highly diffusive and that they might not have experienced. There is also a moral problem in attempting to determine the value of a "social good" whose cost will be apportioned among recipients according to how highly they value it. See Musgrave (1959: 73-89).

5. It is not clear that the political variables that dominate the acquisition of weapons systems have been overcome by the analysts as is evidenced by the typical overruns on costs and underruns on performance of such recent purchases as the F111 and C5A aircraft. In this respect the cost-effectiveness studies of the Pentagon may be primarily academic in value.

6. No attempt is made here to explain basic accounting procedures. There are many readily available sources on this subject, for example, Anthony (1964), Horngren (1967), and Moore and Jaedicke (1967).

7. A discussion of the treatment of depreciation in cost accounting is found in Anthony (1964: 154-172).

8. The underestimation of costs is far more frequent than their overestimation. Certainly a part of this phenomenon is due to the tendency of analysts to be "optimistic" in their cost estimates of projects they favor. For some other reasons see Merewitz (1973).

9. One subsidy we have not mentioned is the set of government services, such as municipal police and fire protection, refuse collection, and so on, provided free of charge to government-sponsored programs. The value of this subsidy might be appropriately estimated from the foregone taxes.

10. The use of experimental techniques for the evaluation of broad-aim programs is beset with difficulties in both conceptualization and application. It has been suggested that a greater emphasis be put on the evaluation of process than end product (Weiss and Rein, 1972; Tribe, 1973).

REFERENCES

Anthony, R. N. *Management Accounting.* Homewood, Ill.: Irwin, 1964.

Arrow, Kenneth. *Social Choice and Individual Values.* New York: Wiley, 1951.

Atkinson, A. B. On the measurement of inequality. *Journal of Economic Theory,* 1970, 2: 221-224.

Azzi, Corry F. and James C. Cox. Equity and efficiency in program evaluation. *Quarterly Journal of Economics,* 1973, 495-502.

Baumol, W. J. On the social rate of discount. *American Economic Review,* 1968, 58: 788-802.

Becker, G. S. *Human Capital.* Princeton: Princeton University Press, 1964.

Block, J. H. *Mastery Learning.* New York: Holt, Rinehart and Winston, 1971.

Chase, Samuel B., Jr. (ed.) *Problems in Public Expenditure Analysis.* Washington, D.C.: Brookings Institution, 1968.

Chipman, J. S. The nature and meaning of equilibrium in economic theory. Pp. 35-64 in D. Martindale (ed.), *Functionalism in the Social Sciences: The Strength and Limits of Functionalism in Anthropology, Economics, Political Science, and Sociology.* Philadelphia: American Academy of Political and Social Science.

Clawson, Marian and Jack L. Knetsch. *Economics of Outdoor Recreation.* Baltimore: Johns Hopkins Press, 1966.

Corazzini, Arthur. The decision to invest in vocational education. *Journal of Human Resources,* 1968, 3: 82-120.

Dorfman, Robert (ed.) *Measuring Benefits of Government Investments.* Washington, D.C.: Brookings Institution, 1965.

Eckstein, Otto. *Water-Resource Development.* Cambridge: Harvard University Press, 1958.

Fein, Rashi. *Economics of Mental Illness.* New York: Basic Books, 1958.

Fishburn, C. *Decision and Value Theory.* New York: Wiley, 1964.

Fisher, G. H. The role of cost-utility analysis in program budgeting. In David Novick (ed.), *Program Budgeting.* Washington, D.C.: U.S. Government Printing Office, 1964.

Forst, B. E. An analysis of alternative periodic health examination strategies. Pp. 393-409 in Niskanen et al. (1973).

Galbraith, J. K. Economics as a system of belief. *American Economic Review,* 1970, 60: 469-478.

Gillespie, W. I. Effect of public expenditures on the distribution of income. Pp. 122-186 in R. A. Musgrave (ed.), *Essays in Fiscal Federalism.* Washington, D.C.: Brookings Institution, 1965.

Gintis, Herbert. Consumer behavior and the concept of soverignity. *American Economic Review,* 1972, 62: 267-278.

Goldman, T. A. (ed.) *Cost-Effectiveness Analysis.* New York: Praeger, 1967.

Goldstein, J. H. The effectiveness of manpower training programs: a review of research on the impact on the poor. Pp. 338-373 in Niskanen et al. (1973).

Hanoch, G. An economic analysis of earnings and schooling. *Journal of Human Resources,* 1967, 2: 310-329.

Haveman, R. H. The ex-post evaluation of navigation improvements. Pp. 249-276 in Niskanen et al. (1973).

Hennes, J. D. The measurement of health. *Medical Care Review,* 1972, 1268-1288.

Hicks, J. R. *Value and Capital.* 2nd ed. London: Oxford University Press, 1946.

Hinrichs, H. H. and G. M. Taylor. *Program Budgeting and Benefit-Cost Analysis.* Pacific Palisades: Goodyear, 1969.

Hirsch, Werner Z. Cost functions of an urban government service: refuse collection. *Review of Economics and Statistics,* 1965, 47: 87-92.

———. *The Economics of State and Local Government.* New York: McGraw-Hill, 1970.

Hitch, C. J. and R. N. McKean. *The Economics of Defense in the Nuclear Age.* Cambridge: Harvard University Press, 1960.

Horngren, C. T. *Cost Accounting.* Englewood Cliffs, N.J.: Prentice-Hall, 1967.

Hu, Tei-wei, Maw Lin Lee and E. W. Stronsdorfer. Economic returns to vocational and comprehensive high school graduates, *Journal of Human Resources,* 1971, 25-50.

Jamison, Dean, D. Fletcher, Pratrick Suppes and Richard Atkinson. Cost and performance of computer-assisted instruction for education of disadvantaged children. Paper presented at the National Bureau of Economic Research Conference on Education as an Industry, Chicago, June 1971.

Kain, J. F. An analysis of metropolitan transportation systems. Pp. 155-187 in Thomas A. Goldman (ed.), *Cost-Effectiveness Analysis.* New York: Praeger, 1967.

Klarman, Herbert E. Syphilis control programs. Pp. 367-414 in Robert Dorfman (ed.), *Measuring Benefits of Government Investments.* Washington, D.C., 1965.

Krutilla, J. V. and Otto Eckstein. *Multiple Purpose River Development.* Baltimore: Johns Hopkins Press, 1958.

Krutilla, J. V. and C. J. Cicchetti. Evaluating benefits of environmental resources with special application to the Hells Canyon. Pp. 447-475 in Niskanen et al. (1973).

Lee, R. D., Jr., and R. W. Johnson. *Public Budgeting Systems.* Baltimore: University Park Press, 1973.

Levin, H. M. A cost-effectiveness analysis of teacher selection. *Journal of Human Resources,* 1970, 5: 24-33. (a)

———. A new model of school effectiveness. Chap. 3 in *Do Teachers Make a Difference?* Washington, D.C.: U.S. Department of Health, Education, and Welfare, 1970. (b)

Lifson, Melvin W. Value Theory. Pp. 79-112 in J. Morley English (ed.), *Cost-Effectiveness—The Economic Evaluation of Engineered Systems.* New York: Wiley, 1968.

Little, I. M. D. *A Critique of Welfare Economics.* 2nd ed. London: Oxford University Press, 1957.

Mack, Ruth P. and Sumner Myers. Outdoor recreation. Pp. 71-116 in Robert Dorfman (ed.), *Measuring Benefits of Government Investments.* Washington, D.C.: Brookings Institution, 1965.

Maidlow, S. T. and Howard Berman. The economics of heroin treatment. *American Journal of Public Health,* 1972, 1397-1406.

Marglin, S. A. The social rate of discount and the optimal rate of investment. *Quarterly Journal of Economics,* 1963, 77: 95-112.

Menz, F. D. Economics of disease prevention: infectious kidney disease. *Inquiry,* 1971, 3-18.

Merewitz, Leonard. Cost Overruns in Public Works. Pp. 227-295 in Niskanen et al. (1973).

Mohring, Herbert. Urban highway investments. Pp. 231-291 in Robert Dorfman (ed.), *Measuring Benefits of Government Investments.* Washington, D.C.: Brookings Institution, 1965.

Moore, C. L. and R. K. Jaedicke. *Managerial Accounting.* Cincinnati: South-Western Publishing, 1967.

Morris, Douglas, and Luther Tweeten. The cost of controlling crime. *Annals of Regional Science of the Western Regional Science Association,* 1971, 5: 33-49.

Musgrave, R. A. Cost-benefit analysis and the theory of public finance. *Journal of Economic Literature,* 1969, 7: 797-806.

———. *The Theory of Public Finance.* New York: McGraw-Hill, 1959.

Musgrave, Richard A. and Peggy B. Musgrave. *Public Finance in Theory and Practice.* New York: McGraw-Hill, 1973.

Mushkin, Selma J. Health as an investment. *Journal of Political Economy,* 1962, 70: 129-142.

Niskanen, W. A., A. C. Harberger, R. H. Haveman, Ralph Turvey and Richard Seckhauser (eds.) *Benefit-Cost and Policy Analysis 1972.* Chicago: Aldine, 1973.

Perl, Lewis J. Family background, secondary school expenditure, and student ability. *Journal of Human Resources,* 1973, 6: 156-180.

Pigou, A. L. *A Study in Public Finance.* 3rd ed. London: Macmillan, 1951.

Prest, A. R. and R. Turvey. Cost-benefit analysis: a survey. *Economic Journal,* 1965, 75: 683-735.

Psacharopoulos, George. *Returns to Education.* San Francisco: Jossey-Bass, 1973.

Quade, E. S. *Analysis for Military Decisions.* Chicago: Rand McNally, 1964.

Ribich, Thomas. *Education and Poverty.* Washington, D.C.: Brookings Institution, 1968.

Rice, D. P. The Direct and Indirect Cost of Illness. Pp. 469-490 in *Federal Programs for the Development of Human Resources.* A compendium of papers submitted to the Subcommittee on Economic Progress, 1968.

Rivlin, Alice. *Systematic Thinking for Social Action.* Washington, D.C.: Brookings Institution, 1971.

Robinson, James and Gerald Smith. The effectiveness of correctional programs. *Journal of Research in Crime and Delinquency.* 1971, 17: 67-80.

Rothenberg, Jerome. *Economic Evaluation of Urban Renewal.* Washington, D.C.: Brookings Institution, 1967.

Samuelson, Paul A. *Economics.* 8th ed. New York: McGraw-Hill, 1970.

Schultz, Theodore W. Capital formation by education. *Journal of Political Economy,* 1960, 68: 571-583.

Schultze, Charles L. *The Politics and Economics of Public Spending.* Washington, D.C.: Brookings Institution, 1969.

Sclowsky, Marcelo. An attempt to estimate rates of return to investment to infant nutrition. Pp. 410-428 in Niskanen et al. (1973).

Smith, K. R., Marianne Miller and F. L. Golladay. An analysis of the optimal of inputs in the production of medical services. *Journal of Human Resources,* 1972, 7: 208-225.

Somers, G. G. and E. W. Stromsdorfer. A cost-effectiveness analysis of in-school and summer

neighborhood youth corps: nationwide analysis. *Journal of Human Resources,* 1972, 7: 446-459.

Suppes, Patric, and Mona Morningstar. Computer-assisted instruction. *Science,* 1969, 166: 343-350. (a)

———. Evaluation of three computer-assisted instruction programs. Psychology Series. Stanford, 1969. (b)

Tribe, L. H. Policy science: analysis or ideology? Pp. 3-47 in Niskanen et al. (1973).

Weisbrod, B. A. *Economics of Public Health.* Philadelphia: University of Pennsylvania Press, 1961.

———. Preventing high school dropouts. Pp. 117-148 in Robert Dorfman (ed.) *Measuring Benefits of Government Investments.* Washington, D.C.: Brookings Institution, 1965.

———. Income redistribution effects and benefit-cost analysis. Pp. 177-209 in Samuel B. Chase, Jr. (ed.), *Problems in Public Expenditures Analysis.* Washington, D.C.: Brookings Institution, 1968.

Weiss, R. S. and Martin Rein. The evaluation of broad-aim programs: difficulties in experimental design and an alternative. Pp. 236-249 in Carol H. Weiss (ed.), *Evaluating Action Programs.* Boston: Allyn and Bacon, 1972.

Wildavsky, Aaron. *The Politics of the Budgetary Process.* Boston: Little, Brown, 1964.

Williams, Alan. Cost-benefit analysis: bastard science? and/or insidious poison in the body politick? Pp. 48-74 in Niskanen et al. (1973).